

Thilini Wijesiriwardene

Ph.D. Candidate, College of Engineering and Computing, University of South Carolina

[🏠 Homepage](#) | [✉️ thilini@sc.edu](mailto:thilini@sc.edu) | [🌐 thilini-w](#) | [🔗 Thiliniw](#) | [🎓 Google Scholar](#)

Columbia, South Carolina - 29201, United States

SUMMARY

I conduct research in Natural Language Processing and Language Modeling, investigating the capabilities of current Large Language Models for analogical reasoning and concept abstraction in the presence of external knowledge.

WORK EXPERIENCE

- **Artificial Intelligence Institute, University of South Carolina** [🌐](#) 08/2019 - Present
Graduate Research Assistant (R.1), Graduate Teaching Assistant and Lab Monitor South Carolina, USA
- **National Library of Medicine** [🌐](#) 05/2021 - 01/2022
Research Intern (R.2) with Dr. Olivier Bodenreider Maryland, USA
- **Wright State University** [🌐](#) 01/2018 - 08/2019
Graduate Research Assistant (R.3) Ohio, USA
- **Wright State University** [🌐](#) 09/2017 - 12/2017
Volunteer Researcher (R.3) Ohio, USA
- **University of Colombo School of Computing** [🌐](#) 03/2015 - 06/2017
Assistant Lecturer Colombo, Sri Lanka
- **University of Colombo School of Computing** [🌐](#) 10/2014 - 03/2015
Graduate Instructor Colombo, Sri Lanka
- **attune Consulting** [🌐](#) 02/2014 - 10/2014
Associate SAP Consultant (Specialized in Sales and Distribution module) Colombo, Sri Lanka

EDUCATION

- **University of South Carolina** 08/2019 - Present
Doctor of Philosophy, Computer Science South Carolina, USA
 - **Advisor:** Dr. Amit Sheth.
 - **Relevant Coursework:** Artificial Intelligence, Reinforcement Learning, Natural Language Processing.
 - **Grade Point Average:** 4.0
- **Wright State University** 01/2018 - 08/2019
MS in Computer Science (Transferred to Ph.D.) Ohio, USA
 - **Advisor:** Dr. Amit Sheth.
 - **Relevant Coursework:** Algorithm Design and Analysis, Deep Learning, Information Retrieval.
 - **Grade Point Average:** 3.89
- **University of Colombo School of Computing** 02/2014
Bachelor of Science in Information and Communication Technology Colombo, Sri Lanka
 - **Relevant Coursework:** Data Mining, , Research Methods.
 - **Grade Point Average:** 3.61

RESEARCH PROJECTS

[R=RESEARCH PROJECT](#)

- [R.1] **Analogy and Abstraction in the Era of Large Language Models** [🌐](#) | *Publications: C.1, C.2, C.3, J.1, W.2*
 - Led a research project exploring the abilities of Large Language Models to perform analogical reasoning and abstraction in the presence of structured knowledge.
 - Published several resources in the form of datasets (See C.3 and C.1).
- [R.2] **Vocabulary Alignment in UMLS Metathesaurus** | *Publications: C.5, W.1, PP.1*
 - Trained a BERT-based language model for synonymy prediction in the UMLS Metathesaurus, achieving an accuracy rate exceeding 90%.
 - Developed and evaluated learning models to perform vocabulary alignment at scale in UMLS Metathesaurus.
- [R.3] **Context-Aware Harassment Detection on Social Media** [🌐](#) | *Publication: C.6*
 - Explored how the meaningful use of context in online interactions can enable highlighting or exonerating tweets with purported toxicity.
 - Introduced a dataset comprising individual tweets paired with their contextual information.

PUBLICATIONS

C=CONFERENCE, J=JOURNAL, W=WORKSHOPS, *=SHARED AUTHORSHIP

- [C.1] Thilini Wijesiriwardene, Ruwan Wickramarachchi, Sreeram Reddy Vennam, Vinija Jain, Aman Chadha, Amitava Das, Ponnurangam Kumaraguru and Amit Sheth. **"KnowledgePrompts: Exploring the Abilities of Large Language Models to Solve Proportional Analogies via Knowledge-Enhanced Prompting."** In Proceedings of the 31st International Conference on Computational Linguistics 2025. [Link](#)
- [C.2] Thilini Wijesiriwardene, Ruwan Wickramarachchi, Aishwarya Naresh Reganti, Vinija Jain, Aman Chadha, Amit Sheth and Amitava Das. **"On the relationship between sentence analogy identification and sentence structure encoding in large language models."** In Findings of the Association for Computational Linguistics: EACL 2024. [Link](#)
- [J.1] Thilini Wijesiriwardene, Amit Sheth, Valerie L. Shalin and Amitava Das. **"Why Do We Need Neurosymbolic AI to Model Pragmatic Analogies?."** In IEEE Intelligent Systems 2023. [Link](#)
- [C.3] Thilini Wijesiriwardene, Ruwan Wickramarachchi, Bimal Gajera, Shreeyash Gowaikar, Chandan Gupta, Aman Chadha, Aishwarya Naresh Reganti, Amit Sheth, and Amitava Das. **"ANALOGICAL - A Novel Benchmark for Long Text Analogy Evaluation in Large Language Models."** In Findings of the Association for Computational Linguistics: ACL 2023. [Link](#)
- [C.4] Riccardo Tommasini, Filip Ilievski, and Thilini Wijesiriwardene. **"IMKG: The internet meme knowledge graph."** In The Semantic Web: 20th International Conference, ESWC 2023. [Link](#)
- [W.1] Goonmeet Bajaj, Vinh Nguyen, Thilini Wijesiriwardene, Hong Yung Yip, Vishesh Javangula, Amit Sheth, Srinivasan Parthasarathy, and Olivier Bodenreider. **"Evaluating Biomedical Word Embeddings for Vocabulary Alignment at Scale in the UMLS Metathesaurus Using Siamese Networks."** In Proceedings of the Third Workshop on Insights from Negative Results in NLP 2022. [Link](#)
- [C.5] Vinh Nguyen, Hong Yung Yip, Goonmeet Bajaj, Thilini Wijesiriwardene, Vishesh Javangula, Srinivasan Parthasarathy, Amit Sheth, and Olivier Bodenreider. **"Context-enriched learning models for aligning biomedical vocabularies at scale in the UMLS Metathesaurus."** In Proceedings of the ACM Web Conference 2022 (WWW '22). Association for Computing Machinery. [Link](#)
- [W.2] Thilini Wijesiriwardene, Ruwan Wickramarachchi, Valerie L. Shalin, Amit P. Sheth. **"Towards efficient scoring of student-generated long-form analogies in STEM."** Workshop on Analogies: from Theory to Applications at ICCBR 2022. [Link](#)
- [C.6] Thilini Wijesiriwardene*, Hale Inan*, Ugur Kursuncu, Manas Gaur, Valerie L. Shalin, Krishnaprasad Thirunarayan, Amit Sheth, and I. Budak Arpinar. **"Alone: A dataset for toxic behavior among adolescents on twitter."** In Social Informatics: 12th International Conference, SocInfo 2020. [Link](#)

OTHER ARTICLES







PP=PREPRINT

- [PP.1] Thilini Wijesiriwardene, Vinh Nguyen, Goonmeet Bajaj, Hong Yung Yip, Vishesh Javangula, Yuqing Mao, Kin Wah Fung, Srinivasan Parthasarathy, Amit P. Sheth, and Olivier Bodenreider. **"Ubert: A novel language model for synonymy prediction at scale in the umls metathesaurus."** arXiv preprint arXiv:2204.12716 (2022). [Link](#)

SKILLS

- **Programming Languages:** Python, C, Java, R, SQL, Prolog.
- **Frameworks and Libraries:** PyTorch, Huggingface, Pandas, Numpy, Matplotlib, Langchain, Scikit-learn.
- **Software & Tools:** GitHub, Anaconda, MS Office, MATLAB, Protege, Neo4j.

TRAVEL GRANTS & OTHER AWARDS

- Student Volunteer: The 31st International Conference on Computational Linguistics 2025, Abu Dhabi, UAE
- Student Volunteer: The 62nd Annual Meeting of the Association for Computational Linguistics 2023, Toronto, Canada 
- Student Scholarship: The 37th AAAI Conference on Artificial Intelligence 2023, Wasginton DC, USA 
- CRA-WP Grad Cohort Workshop for Women 2022, New Orleans, Louisiana, USA 
- Grace Hopper Celebration 2020, Virtual 
- Travel Award: WSTNET Web Science Summer School 2018, Hannover, Germany 
- Travel Award: Tri-State Celebration of Women in Computing (TRIWIC) 2018, Cincinnati, OH, USA 

ACADEMIC SERVICE

R=REVIEWER, AC=AREA CHAIR, PC=PROGRAM COMMITTEE MEMBER

- [R] ACL Rolling Review 2024 (April, June, August, October, December),
Empirical Methods in Natural Language Processing (EMNLP) 2023,
ACM SIGCHI Conference on Computer-Supported Cooperative Work & Social Computing (CSCW) 2022,
ACM Transactions on Computing for Healthcare 2024.
- [AC] ACL Rolling Review (2025 February).
- [PC] Extended Semantic Web Conference (ESWC) - Resource Track 2025.